

Privacy-preserving Framework for Automated Detection of Arrhythmia in ECG Data

Kacper Gil and Andres Vejar

AGH University of Krakow, Kraków, Poland

<https://doi.org/10.26636/jtit.2025.FITCE2024.2042>

Abstract — The integration of machine learning in biomedical engineering applications is crucial to ensure user data security and privacy. This work explores anonymization and differential privacy (DP) frameworks to reduce the risk of biometric identification. The DP method is used to train models in biosignal data without compromising the diagnostic results. The proposed approach for privacy-preserving arrhythmia detection uses a machine learning diagnostic system that reduces discrepancies between preprocessed and raw data, maintaining a correct level of diagnostic precision while improving privacy. The application is evaluated using a control model to analyze the accuracy difference when using privacy-preserving input data.

Keywords — arrhythmia detection, differential privacy, ECG data, privacy enhancing technologies

1. Introduction

Automated diagnostic systems allow reducing the load on health facilities and contribute to improving the quality of medical care at home. Such systems require tracking of several biosignals to monitor the health status of patients. It is important to consider privacy-enhanced methods in the diagnostic system, given that these signals, like electroencephalogram (EEG) and or electrocardiogram (ECG) can reveal the identities of patients using biometric identification methods.

An ideal feature of an automated diagnostic system is the ability to ensure privacy by design [1], [2], where privacy should be built into the technology that supports the system. Important elements to consider during the design phase are the minimization of the user data, the controllability of personal data, the transparency about the system operation, the control on which authorized entities can have data access, and the secure segregation of the data.

1.1. Privacy-enhancing Technologies

For practical engineering implementations, several privacy-enhancing technologies (PET) are available in the literature. Paper [3], specified three general categories: algorithmic PETs, where a formal definition of the algorithms allow to specify strict privacy requirements, architectural PETs, where privacy is enhanced by the design of the underlying distributed computation system, and augmentation PETs, where improve-

Tab. 1. Categories of privacy-enhancing technologies (PETs).

Algorithmic PETs	Architectural PETs	Augmentation PETs
Differential privacy [4]	Federated learning [5]	Synthetic data [6]
Zero-knowledge proofs [7]	Multi-party computation [8]	Digital twinning [9]
Homomorphic encryption [10]		

ment of the user privacy by the incorporation of generative models of synthetic data and digital twinning is explored. These categories, and relevant examples are presented in Tab. 1. For algorithmic PETs, the most important examples are differential privacy (DP), zero-knowledge proofs, and homomorphic encryption.

If an external observer cannot verify that the information of a particular user was involved in the computation, then the algorithm is differentially-private. A similar concept concerns zero-knowledge proofs, where two parties, the *verifier* and the *prover* interact to acknowledge the possession of information. The *prover* goal is to acknowledge information possession without disclosing it. Another type of algorithmic PET is homomorphic encryption, it considers the ability of a cryptosystem to perform computations in encrypted data. Upon decryption, it yields an output that is exactly the same as if the operations had been carried out on the unencrypted data.

For architectural PETs, federated learning consist on a distributed strategy, where machine learning models are trained locally, and only the parameters of the models are communicated between the federated peers. Multi party-computation refers to the use of private data in protected computation tasks. All the parties can have access the computed results, but the computation will not reveal the individual data to the peers.

The last type of PETs, refers to augmentation. Synthetic data is data that was created to support and test algorithms and mathematical models, it is specially important in data science and machine learning tasks. A specific type of augmentation is the digital twinning, where a virtual counterpart of a physical system is created to study the real system and to predict its

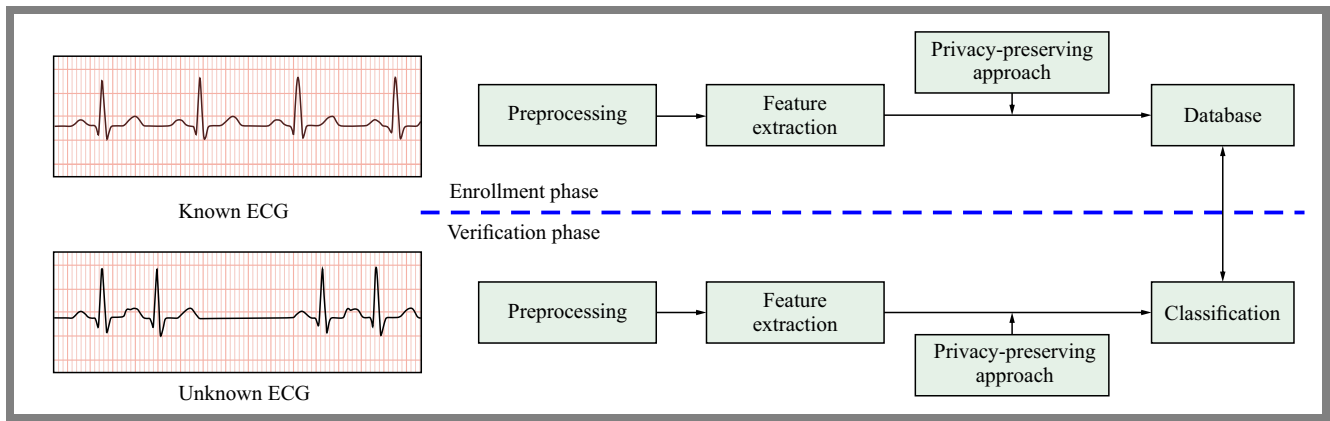


Fig. 1. Biometric identification and re-identification using ECG signals.

response given artificial stimuli. For example, digital twins can be customized to study the evolution of a medical therapy and to predict future results. Anonymized twins can be used by third party healthcare contractors in order to understand the problem under study and to propose privacy-preserving systems to the healthcare facilities.

This work focuses on algorithmic PETs, by the use of DP methods in automated detection of arrhythmia.

1.2. Biometric Identification

There are two phases to the process of biometric identification, namely the enrollment and the verification phases. The enrollment phase is the process of registering a source of biometric data jointly with its associated identification index, with the possibility of including other diverse biometric data, e.g. fingerprints and face image. The data stored are generally processed to obtain a set of features that are characteristic to one person, the biometric template data. The verification phase consists of matching the template data into new data. This phase can be challenging, because biometric data can vary from measurement to measurement.

Biometric identification and authentication using ECG [11]–[13] can be achieved directly or in conjunction with other sources of biometric data. It is interesting because it can be used as a continuous authentication method in critical systems, for example in continuous driver authentication for cash transport, public transportation, military, and car rental and sharing services [14].

The working principle of biometric identification of ECG in a diagnostic system can be seen in the Fig. 1. The enrollment phase consist on the preprocessing of the *Known ECG* signal, i.e. creating a pair (ECG signal, user ID). In order to compare the ECG signal with another *Unknown ECG* signal it is required to extract features of the signal. These features will be stored as a template in the system database. A privacy-preserving approach for registering the ECG features in the database will enhance the security diagnostic service. For example, only enrolled users will be able to be classified by the arrhythmia detection service. In this work, we focus in implementing a differentially-private classification model for

ECG diagnostic, guarantying that no database sample can significantly affect the outcome of the classification.

1.3. Medical Background of Arrhythmia

Arrhythmia is a medical condition characterized by an irregular heartbeat, also classified as tachycardia or bradycardia if the heart beats too fast or too slow, respectively. Alternatively, the irregularity can display no pattern; in such cases it is called fibrillation. Factors of increased risk of arrhythmia include cardiovascular disease, heart surgery, and cardiomyopathy that implies changes in heart structure. Other causes not related to the heart are electrolyte imbalances, medications, and certain stimulants. Personal lifestyle also plays a role in the incidence rate of heart irregularities. High levels of stress, smoking, and physical exertion are the most common. Generally, arrhythmias manifest only as palpitations, light dizziness, and shortness of breath. However, in more severe cases it can lead to fainting and may even be life-threatening. The diagnosis procedure involves the use of an ECG, usually taken over a period of 24–48 h, with the help of a Holter monitor.

Several arrhythmia detection methods [15]–[17], can be found in the literature, where the Physionet computing in cardiology challenge and its ECG dataset is an important benchmark for machine learning methods [18]. More advanced data sets are also available, for example, the 12 leads ECG data set [19].

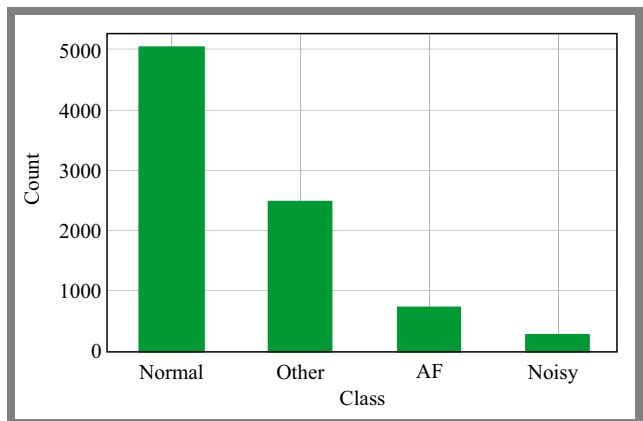


Fig. 2. Distribution of target classes in the dataset.

Therefore, there are a growing collection of automated methods of arrhythmia detection and classification [20] that can benefit greatly with the incorporation of PETs.

2. Materials and Methods

2.1. Data Sourcing and Labeling

The study from [21] proposes feature-based classifiers and convolutional neural network (CNN) models for arrhythmia classification using the first minute of each ECG sample. In this work, we use a similar approach considering a CNN model, but we generate 2D images of the ECG samples using recurrence plots [22]. Another important difference is that we restrict the length of input data for the CNN classifier to a very short time period of around 2.3 s. To train and to test the model, we use random sampling to consider any period of 2.3 s in the full length timeseries. Our model is designed for real time detection using the buffer of 2.3 s.

The data was provided by AliveCor for the purposes of the aforementioned challenge [18]. The total number of ECG recordings exceeded 12 000. Each of them was taken using single-channel ECG devices of one of three generations. The electrodes, mostly, were placed in each hand of a patient, resulting in lead I (LA-RA) ECG. Many of the data series were inverted, creating (RA-LA) series. The signal recordings average at around 30 s. The equipment then transmitted the data to a portable device over radio waves using 19 kHz carrier frequency and a modulation index of 200 Hz/mV. The data was digitized in 16-bit files with a sample frequency of 300 Hz.

The experts have divided the data into 4 classes: 0 – normal rhythm, 1 – atrial fibrillation (AF) rhythm, 2 – other (abnormal) rhythm, and 3 – noisy recording. The distribution of the classes can be seen in the Fig. 2.

2.2. System Description

In Fig. 3 the diagram of the proposed approach to preserving the privacy of the arrhythmia detection system is presented. This research examines a machine learning diagnostic system in which raw ECG biosignals x undergo client-side pre-processing to become a filtered signal u . Subsequently, this signal is utilized by the diagnostic system g at the diagnostic center. The goal of this system is to reduce the discrepancy between the results of the preprocessed g and a raw data classifier f , $f(x) \approx g(u)$, thus maintaining high diagnostic precision while improving privacy. The application is tested with the control model f that is not privacy preserving, to compare the accuracy level of the arrhythmia detection.

The use of recurrence plots and phase space analyses have seen use in some classification approaches [23], [24].

Out of the input data 700-samples-long snippets (2.3 s) were randomly cropped and later transformed into image data using the `RecurrencePlot` function from the `pyts` library [25]. The threshold and percentage values were set to “point” and 20 respectively. These parameters are used for binarization

of the recurrence plot, that consist in a 700×700 pixel image with a single color channel. The images were resized using bilinear interpolation to 350×350 pixels.

This data was then shuffled and passed to the CNN model consisting in three 2D-convolution layers for acquiring image features and three dense layers acting as the output classifier. The total number of trainable parameters of the CNN model is 9 539 669.

In this application we aggregate the target classes into two: 0 – normal rhythm and 1 – atrial fibrillation (AF) rhythm or other (abnormal) rhythm. Noisy recordings are excluded from the dataset given that, this is a problem that needs to be addressed early, during signal acquisition [26], [27].

The privacy levels are controlled by the parameters for $\epsilon > 0$ and $\delta \in [0, 1)$.

The classifier g , that takes an input u and returns the output y , is (ϵ, δ) -differentially-private for two similar datasets U_a and U_b , $U_a \cap U_b \neq \emptyset$ if the following relation is established:

$$\mathbf{P}(g(u_a \in U_a)) \leq \exp(\epsilon) \mathbf{P}(g(u_b \in U_b)) + \delta. \quad (1)$$

Smaller choices of the parameter ϵ make the model more private, controlling the level of noise. The parameter δ refers to the probability of a data breach. It is pertinent to set ϵ and δ to achieve a trade-off between privacy and classification performance.

3. Results

In order to compare a privacy-unaware classification model of arrhythmias with a differentially private model, we trained a baseline model and a differentially-private model. Both models share the same CNN architecture. The DP training of the models was performed using the `Opacus` library [28]. The models were developed using the `PyTorch` library for deep learning in a GPU NVIDIA GeForce RTX 4080.

3.1. Baseline CNN Classification

The initial output of the CNN classification model, constructed to explore the performance of DP models, can be seen in the Fig. 4.

One important application of online detection and diagnostic systems is to trigger alarms or alert the corresponding medical services in case of an improper heart rhythm. Consequently, the most important factor to minimize is the number of false negatives in the classification. We use the false omission

Tab. 2. Model metrics.

Metric	Initial model	Final model
FOR	0.3982	0.1274
Accuracy	0.7846	0.8846
Precision	0.8608	0.9029
Sensitivity	0.6018	0.8230
Specificity	0.9252	0.9328
F1 score	0.7083	0.8611

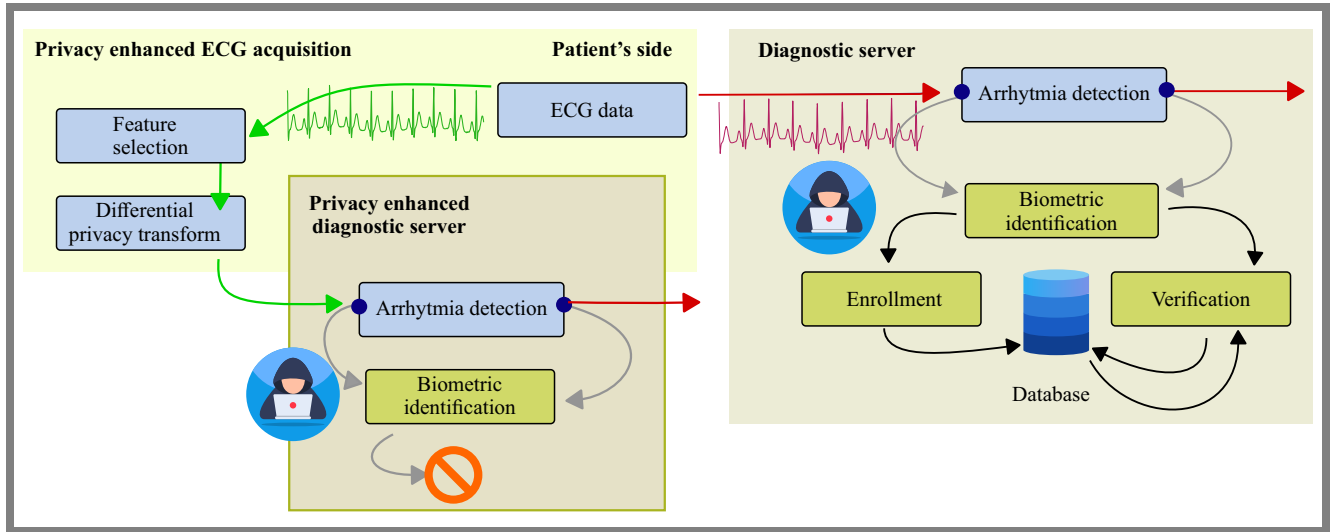


Fig. 3. System diagram considering raw and privacy-enhanced arrhythmia detection. In the left side is presented the proposed privacy enhanced arrhythmia diagnostic system. A compromised, raw ECG arrhythmia diagnostic system is depicted in the right side.

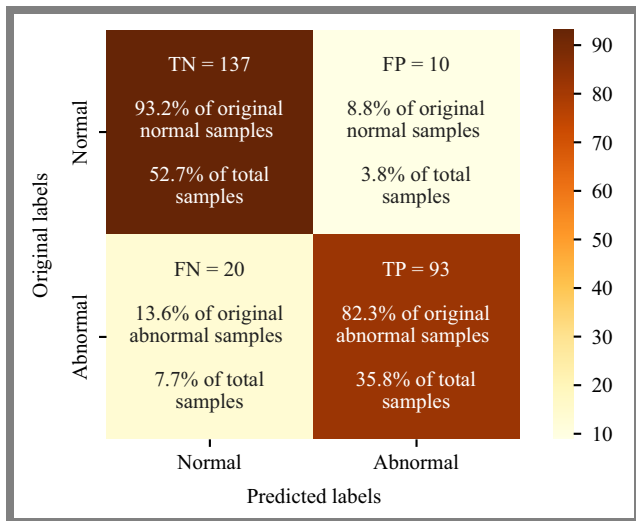


Fig. 4. Confusion matrix of the baseline model.

rate (FOR), to measure the proportion of incorrect negative classifications, false negatives (FN) with respect to the overall negative class:

$$FOR = \frac{FN}{TN + FN} \tag{2}$$

The selected metrics for comparison of the baseline model in the initial and final versions are presented in Tab. 2.

The metrics presented in Tab. 2 show some notable improvements after optimization, including a reduction in the false omission rate (from 0.3982 to 0.1274), an increase in sensitivity (from 0.6018 to 0.8230), and an improved F1 score (from 0.7083 to 0.8611). These improvements suggest that the optimization efforts have enhanced the classifier’s ability to detect cardiac arrhythmias, particularly in minimizing missed detections. Said parameter is of main concern as it can be feasibly presumed that the potential user will already have a history of prior medical issues with heart rhythm. That is, a possible false alarm will not be as damaging as a missed anomaly, given that the user will have a way of turning it off.

The remaining parameters also saw an increase, i.e. precision (from 0.8608 to 0.9029) and specificity (from 0.9252 to 0.9328), which points to the overall improvement of the model.

The final accuracy of 88.46% is a considerable improvement over the initial 78.46%. The model was designed to process short ECG signals as inputs (about 2.3 s), which makes it particularly useful in real-world scenarios where rapid detection is critical. Furthermore, the model is currently in its preliminary stages of development, and future iterations are planned to incorporate greater code complexity, which is expected to further improve performance across all metrics.

3.2. DP Implementation in the CNN

To compare the selection of privacy hyper-parameters with respect to the performance in retraining, firstly a DP version of CNN classifier studied in the previous subsection was trained from scratch with weak privacy parameters until achieving a classification accuracy of 75%. Said threshold was pre-selected for testing the capabilities of DP implementation in ECG classification. It can be potentially improved with training-time optimization techniques.

Secondly, the classifier was loaded and retrained with different choices of the privacy hyper-parameters:

- Maximum gradient norm G – corresponds to the maximum achievable norm for each gradient sample. Greater gradients will be clipped to the value of this parameter. With higher values of the maximum gradient norm higher levels of privacy are achieved.
- Privacy budget P_ϵ – cumulative value of the ϵ parameter over all epochs during training. With smaller ϵ values, higher levels of privacy are achieved.
- δ – the likelihood of a data breach.

For all the experiments, δ was fixed in 1.1, the values of P_ϵ include 12, 48, 84, and 120. The results are presented in Tab. 3 according to maximum gradient norm. Training stage con-

sisted of 20 epochs. The preliminary DP implementation was able to achieve 75% accuracy. Taking that into consideration, it can be observed that higher levels of G values substantially affected the model performance.

4. Conclusions and Further Work

In this work, a privacy-preserving framework for the detection of arrhythmia is presented. The framework considers a *privacy enhanced ECG acquisition* on the patient’s side, useful for remote diagnostic in homecare, and a *privacy enhanced diagnostic server* that provides the automated diagnostic service.

The ongoing work considers a validation of the results with standard ECG biosignal databases. It is important to evaluate the performance penalty of implementing DP, or other PETs in standard deep learning models. One of the objectives of this work is to promote the application of privacy-enhancing technologies in the early stages of automated diagnostic systems, revisiting well proven classification methods and incorporating privacy-enhancing hyper-parameters during design and learning phases.

Further work will consider the design of automated diagnostic systems with the joint goal of security and privacy. In addition, the use of a large set of sensors (e.g. temperature, pulse oximetry, EEG, and EMG) and different target diseases for detection can be explored as an extension of the proposed approach.

Tab. 3. Performance after DP retraining.

G	P_ϵ	Train loss	Train acc.	Test loss	Test acc.
1.10	120	1.03	73.35	1.01	72.99
1.10	84	1.03	73.37	1.00	73.00
1.10	48	1.04	73.23	1.01	73.19
1.10	12	1.06	72.89	1.02	72.45
4.07	120	0.91	50.70	0.70	49.57
4.07	84	0.97	50.33	0.70	49.51
4.07	48	0.83	50.02	0.70	49.56
4.07	12	4.45	50.55	1.06	51.30
7.03	120	2.25	50.31	0.79	51.18
7.03	84	4.99	50.74	3.76	51.52
7.03	48	21.45	50.35	16.60	51.03
7.03	12	65.63	50.62	188.09	50.65
10.00	120	21.18	50.24	3.48	50.40
10.00	84	45.87	50.80	0.98	50.58
10.00	48	7.30	50.29	1.02	51.54
10.00	12	168.60	50.35	50.34	50.05

Acknowledgments

This research was supported by the National Research Institute, grant number POIR.04.02.00-00-D008/20-01, on “National Laboratory for Advanced 5G Research” (acronym PL-5G) as part of the Measure 4.2 Development of modern research infrastructure of the science sector 2014–2020 financed by the European Regional Development Fund.

References

- [1] P. Schaar, “Privacy by Design”, *Identity in the Information Society*, vol. 3, pp. 267–274, 2010 (<https://doi.org/10.1007/s12394-010-0055-x>).
- [2] A. Nordgren, “Privacy by Design in Personal Health Monitoring”, *Health Care Analysis*, vol. 23, pp. 148–164, 2013 (<https://doi.org/10.1007/s10728-013-0262-3>).
- [3] S. Jordan, C. Fontaine, and R. Hendricks-Sturup, “Selecting Privacy-enhancing Technologies for Managing Health Data Use”, *Frontiers in Public Health*, vol. 10, 2022 (<https://doi.org/10.3389/fpubh.2022.814163>).
- [4] M. Yang *et al.*, “Local Differential Privacy and its Applications: A Comprehensive Survey”, *Computer Standards & Interfaces*, vol. 89, art. no. 103827, 2023 (<https://doi.org/10.1016/j.csi.2023.103827>).
- [5] K.K. Coelho *et al.*, “A Survey on Federated Learning for Security and Privacy in Healthcare Applications”, *Computer Communications*, vol. 207, pp. 113–127, 2023 (<https://doi.org/10.1016/j.comcom.2023.05.012>).
- [6] M. Giuffrè and D.L. Shung, “Harnessing the power of synthetic data in healthcare: innovation, application, and privacy”, *Digital Medicine*, vol. 6, art. no. 186, 2023 (<https://doi.org/10.1038/s41746-023-00927-3>).
- [7] L. Petrosino *et al.*, “A Zero-knowledge Proof Federated Learning on DLT for Healthcare Data”, *Journal of Parallel and Distributed Computing*, vol. 196, art. no. 104992, 2024 (<https://doi.org/10.1016/j.jpdc.2024.104992>).
- [8] T. Liu, “Research on Privacy Techniques Based on Multi-Party Secure Computation”, *2024 3rd International Conference on Artificial Intelligence and Autonomous Robot Systems (AIARS)*, Bristol, United Kingdom, 2024 (<https://doi.org/10.1109/AIARS63200.2024.00171>).
- [9] C.S. Jørgensen, A. Shukla, and B. Katt, “Digital Twins in Healthcare: Security, Privacy, Trust and Safety Challenges”, *Proc. of European Symposium on Research in Computer Security*, pp. 140–153, 2023 (https://doi.org/10.1007/978-3-031-54129-2_9).
- [10] K. Munjal and R. Bhatia, “A Systematic Review of Homomorphic Encryption and its Contributions in Healthcare Industry”, *Complex & Intelligent Systems*, vol. 9, pp. 3759–3786, 2023 (<https://doi.org/10.1007/s40747-022-00756-z>).
- [11] A.D.C. Chan, M.M. Hamdy, A. Badre, and V. Badee, “Person Identification using Electrocardiograms”, *2006 Canadian Conference on Electrical and Computer Engineering*, Ottawa, Canada, 2006 (<https://doi.org/10.1109/CCECE.2006.277291>).
- [12] J. Xu, T. Li, Y. Chen, and W. Chen, “Personal Identification by Convolutional Neural Network with ECG Signal”, *2018 International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju, South Korea, 2018 (<https://doi.org/10.1109/ICTC.2018.8539632>).
- [13] S. Asadianfam, M.J. Talebi, and E. Nikougoftar, “ECG-based Authentication Systems: A Comprehensive and Systematic Review”, *Multimedia Tools and Applications*, vol. 83, pp. 27647–27701, 2023 (<https://doi.org/10.1007/s11042-023-16506-3>).
- [14] L.D. Chhibbar *et al.*, “Enhancing Security Through Continuous Biometric Authentication Using Wearable Sensors”, *Internet of Things*, vol. 28, art. no. 101374, 2024 (<https://doi.org/10.1016/j.iot.2024.101374>).

- [15] A.Y. Hannun *et al.*, “Cardiologist-level Arrhythmia Detection and Classification in Ambulatory Electrocardiograms Using a Deep Neural Network”, *Nature Medicine*, vol. 25, pp.65–69, 2019 (<https://doi.org/10.1038/s41591-019-0359-9>).
- [16] R. Li *et al.*, “Interpretability Analysis of Heartbeat Classification Based on Heartbeat Activity’s Global Sequence Features and BiLSTM-attention Neural Network”, *IEEE Access*, vol. 7, pp. 109870–109883, 2019 (<https://doi.org/10.1109/ACCESS.2019.2933473>).
- [17] S. Mousavi and F. Afghah, “Inter- and Intra-patient ECG Heartbeat Classification for Arrhythmia Detection: A Sequence-to-Sequence Deep Learning Approach”, *ICASSP 2019–2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, United Kingdom, 2019 (<https://doi.org/10.1109/ICASSP.2019.8683140>).
- [18] G.D. Clifford *et al.*, “AF Classification from a Short Single Lead ECG Recording: the PhysioNet Computing in Cardiology Challenge 2017”, *2017 Computing in Cardiology Conference (CinC)*, Rennes, France, 2017 (<https://doi.org/10.22489/CinC.2017.065-469>).
- [19] E.A. Perez Alday *et al.*, “Classification of 12-lead ECGs: the PhysioNet/Computing in Cardiology Challenge 2020”, *Physiological Measurement*, vol. 41, art. no. 124003, 2020 (<https://doi.org/10.1088/1361-6579/abc960>).
- [20] Y. Ansari, O. Mourad, K. Qaraqe, and E. Serpedin, “Deep Learning for ECG Arrhythmia Detection and Classification: An Overview of Progress for Period 2017–2023”, *Frontiers in Physiology*, vol. 14, art. no. 1246746, 2023 (<https://doi.org/10.3389/fphys.2023.1246746>).
- [21] F. Andreotti *et al.*, “Comparing Feature-based Classifiers and Convolutional Neural Networks to Detect Arrhythmia from Short Segments of ECG”, *2017 Computing in Cardiology Conference (CinC)*, Rennes, France, 2017 (<https://doi.org/10.22489/CinC.2017.360-239>).
- [22] B.M. Mathunjwa *et al.*, “ECG arrhythmia classification by using a recurrence plot and convolutional neural network”, *Biomedical Signal Processing and Control*, vol. 64, art. no. 102262, 2021 (<https://doi.org/10.1016/j.bspc.2020.102262>).
- [23] S.-C. Fang and H.-L. Chan, “QRS Detection-free Electrocardiogram Biometrics in The Reconstructed Phase Space”, *Pattern Recognition Letters*, vol. 34, pp. 595–602, 2013 (<https://doi.org/10.1016/j.patrec.2012.11.005>).
- [24] B. M. Mathunjwa *et al.*, “ECG Recurrence Plot-based Arrhythmia Classification Using Two-dimensional Deep Residual CNN Features”, *Sensors*, vol. 22, art. no. 1660, 2022 (<https://doi.org/10.3390/s22041660>).
- [25] J. Faouzi and H. Janati, “pyts: A Python Package for Time Series Classification”, *Journal of Machine Learning Research*, vol. 21, 2020 (<https://jmlr.csail.mit.edu/papers/volume21/19-763/19-763.pdf>).
- [26] S.E. Mathe, N.K. Penjarla, S. Vappangi, and H.K. Kondaveeti. “Advancements in Noise Reduction Techniques in ECG Signals: A Review”, *2024 IEEE 3rd World Conference on Applied Intelligence and Computing (AIC)*, Gwalior, India, 2024 (<https://doi.org/10.1109/AIC61668.2024.10730852>).
- [27] F. Liu, Y. Xu, and Y. Yao, “Highly Efficient Low Noise Solutions in ECG Signals”, *Journal of Physics: Conference Series*, vol. 2246, art. no. 012030, 2022 (<https://doi.org/10.1088/1742-6596/2246/1/012030>).
- [28] A. Yousefpour *et al.*, “Opacus: User-friendly Differential Privacy Library in PyTorch”, *ArXiv*, 2021 (<https://doi.org/10.48550/arXiv.2109.12298>).

Kacper Gil, B.Eng.

Institute of Telecommunications

 <https://orcid.org/0009-0002-9404-9720>

E-mail: kagil@student.agh.edu.pl

AGH University of Krakow, Kraków, Poland

<https://www.agh.edu.pl>

Andres Vejar, Ph.D.

Institute of Telecommunications

 <https://orcid.org/0000-0002-2041-0387>

E-mail: avejar@agh.edu.pl

AGH University of Krakow, Kraków, Poland

<https://www.agh.edu.pl>